Naval Research Laboratory
Washington, DC 20375-5320

NRL/FR/5513--00-9963

Design and Evaluation of Digital METOC Documents to Support Retrieval and Use of Information: User-Centric CDROM Compared to an Equivalent Paper Document and its Republished Web Version

James A. Ballas
William S. McBride

Navy Center for Applied Research in Artificial Intelligence
Information Technology Division

William C. Kooiman
Robert T. Miyamoto

Applied Physics Laboratory
University of Washington
Seattle, Washington

October 6, 2000

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave Blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 6, 2000	Final October 1, 1997 to September 30, 1999

4. TITLE AND SUBTITLE

Design and Evaluation of Digital METOC Documents to Support Retrieval and Use of Information: User-Centric CDROM Compared to an Equivalent Paper Document and its Republished Web Version

5. FUNDING NUMBERS

PN - 55-7188
PE - 0602233N
TA - 03328

6. AUTHOR(S)

James A. Ballas, William C. Kooiman,* Robert T. Miyamoto,* and William S. McBride

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/FR/5513--00-9963

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

* Applied Physics Laboratory
University of Washington
Seattle, Washington 98105-6698

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution unlimited.	

13. ABSTRACT (*Maximum 200 words*)

A field study was conducted to evaluate how U.S. Navy Meteorological and Oceanographic (METOC) personnel would use a user-centric CDROM (UC-CD), the equivalent paper document (PAPER), and its republished Web version (RP-WEB) to complete several problems that represented their actual work (e.g., prepare weather briefs, retrieve specific METOC facts). A methodology was developed and used to log the time spent in subtasks (i.e., search for information, interpret information) on these problems. Military METOC personnel at two facilities were recruited as subjects. Results showed that the type of interface had general effects on performance. For example, the browse time for information was longer in three problems using the RP-WEB interface compared to using the PAPER document. The longer browse times were not due to Web downloading delays. Recommendations are provided for improved METOC interface design.

14. SUBJECT TERMS		15. NUMBER OF PAGES
HCI	Web interface design	30
METOC	User-centered design	16. PRICE CODE
Interactive CDROM		

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18
298-102

CONTENTS

DESIGN AND EVALUATION OF DIGITAL METOC DOCUMENTS TO SUPPORT RETRIEVAL AND USE OF INFORMATION: USER-CENTRIC CDROM COMPARED TO AN EQUIVALENT PAPER DOCUMENT AND ITS REPUBLISHED WEB VERSION

1. OBJECTIVE

The objective of this research was to observe and compare the use of a traditional paper document (PAPER) to the use of two computer-based documents: a user-centric CDROM (UC-CD) developed using instructional design techniques; and a hyperdocument prepared for the Web (RP-WEB) by republishing the traditional PAPER document. All documents were used to retrieve meteorological and oceanographic (METOC) information to incorporate into a briefing report. This research will assist the Navy in adopting informational technologies by providing an understanding of techniques for designing computer interfaces. The three documents, two computer-based and one paper, are all currently used at Navy METOC facilities in normal operations. One of the computer documents is distributed on CDROM. The other is hosted over the Web. The UC-CD was developed for usability following user-centered or instructional design techniques. The primary goal of the RP-WEB document was to quickly make the information in the PAPER document available via computer. This experiment involved a collaboration between the Naval Research Laboratory (NRL) and the Applied Physics Laboratory at the University of Washington (APL-UW).

2. BACKGROUND

2.1 METOC Operations and Resources

METOC support for the Navy takes many forms. One of the main units is a group that is led by officers, with experienced noncommissioned officers and specialized enlisted personnel. A METOC unit generally provides meteorological and oceanographic overviews and forecasts to the aircraft, ships, and land-based components of the Navy. Most of the units provide weather briefings and other METOC support for pilots and are generally located on or near naval air stations or ships deploying aircraft. The METOC unit provides this information through direct briefings, phone conversations, video, and the Web. Other METOC officers support ships. They often ride aboard the ship without the benefit of other METOC personnel to consult. They prepare weather forecasts and oceanographic summaries. There may be only one officer for a battle group and only when the group is at station. So a typical tour of duty may result in the officer riding on one ship for several weeks and then being transported to another ship for several weeks and so on as the ships enter and leave active operations. Regardless of who they are supporting, the METOC personnel are called on frequently to prepare briefs that describe the impact of METOC information on naval operations. The average number of briefs per year was estimated to be 144.

The resources used by the METOC personnel for briefings are primarily historic data from the Naval Oceanographic Office (NAVOCEANO) based at Stennis Space Center, Mississippi, and forecasting from Fleet Numerical Meteorology and Oceanography Center (FNMOC). They have the latitude to use any source they deem appropriate, such as commercial weather, to validate their official Navy forecasts. This study focuses on some of the products available from NAVOCEANO.

NAVOCEANO has been at the forefront of providing information digitally. In the early 1990s they recognized the need to transition from printed documents to digital documents distributed on CDROM or

Manuscript approved May 17, 2000.

hosted over the Web. One of the challenges of designing digital documents is to overcome its inherently lesser resolution and portability when compared to paper. The digital documents used in this study were created at that time when the digital technologies were in their infancy. The UC-CD document was begun before the Web was widely known. It was created following instructional design techniques borrowed from computer-based training and relied heavily on multimedia technologies. The RP-WEB document was created later, but at a time when the Web was limited to mainly hypertext and static graphics. There has been a steady increase in capabilities of Web browsers since that time to where virtually anything that could be done on CDROM can now be done in a Web document. Although we refer to the documents in this study by the media on which they are hosted, it is the interface differences brought about by the media for which they were developed and the differences in usage due to the different "look and feel" of the documents that we analyze here. Furthermore, with the convergence of technologies, the results that we present can be applied to either Web or CDROM or even Web documents distributed on CDROM.

2.2 User-Centric Hypermedia Design

The military community as a whole, and the METOC community in particular, has been rapidly adopting information technology to meet current mission requirements. One of the fundamental goals of IT-21, the Navy program to introduce information technology (IT), is to improve the speed of command. From the user's perspective, speed of command will depend on how rapidly the information can be retrieved and analyzed. This in turn depends on the design of the information media. Hypermedia, particularly Web page retrieval and delivery, is the primary software technology being used to serve the users. Thus its design is a critical factor in achieving the Navy's goals for IT-21. Unfortunately, effective hypermedia design is a difficult goal to achieve.

In fact, Landauer (1995) showed that productivity gains in information industries did not reflect the degree of investment in information technology. One of the major goals of this technology investment has been to reduce the amount of paper in offices—to achieve the elusive paperless office. However, Landauer concludes that the paperless office has not been achieved for one key reason:

> "Electronic documents would have long since arrived if it were not for one stumbling block: people hate them. In almost every trial of providing electronic documentation, users have found the documents so unpleasant to use and so unsupportive of the work that they are needed for that they have refused to surrender their paper."

We found that there are exceptions to this pessimistic outcome, and these generally occur when the media has been produced through a process called user-centered design.

There are three activities in the user-centric process: analysis, evaluation, and testing. The first refers to an analysis of the user's task and job, and how the information media can support the requirements of the job. The second refers to an iterative evaluation of the media during the development phase. Often this evaluation will involve the intended users. The last refers to a formal testing to determine how well the media will work in actual operations. The development (Miyamoto et al. 1991) of the Digital METOC and Acoustic Reference System (DMARS), the UC-CD used in this study, was based upon the principles of instructional design, which include these three activities. One of the outcomes of the first activity, analyzing the job requirements, was to design the UC-CD to support the preparation of METOC or weather briefings. Support for this task included digital cut/copy operations to move information from the UC-CD into a briefing document. Iterative design activities resulted in abandoning the original strategy of using icons for navigation. It became apparent that images loaded too slowly from the CDROM media for effective use. Ultimately, a text-based navigation menu and new interaction methods were developed to achieve an effective and consistent design. This report is an example the last step: formal testing.

A recent analysis of Air Force weather forecasters illustrates a formal approach to the first activity: task analysis. Pliske et al. (1997) conducted a Cognitive Task Analysis (CTA) of U.S. Air Force forecasters to advance our understanding of the weather forecasting process. CTA typically involves identifying the processes and information that are used by domain experts in making key decisions. Often an analysis of critical incidents is involved. Pliske et al. (1997) developed a general model or framework that included three steps: 1) diagnosing the present weather situation; 2) understanding what is causing the present weather situation; and 3) predicting the future. They also concluded that the highly skilled forecasters were more likely to use a visual mental model in the second step. Potentially, forecasting tools that support the development of a visual mental model of the weather situation will have an important positive impact.

One of the enormous challenges that the Navy faces in moving to IT-21 is replacing the legacy documents that the Fleet uses with computer-based media. In particular, the rapid development of Web sites to distribute information is raising questions about how to validate the information. It also raises an issue at the core of this report: do legacy paper documents serve a function that cannot be met by hypermedia? These issues exist in the commercial sector as well. In a study of the hypermedia publishing industry, Bellotti and Rogers (1997) examined workflow processes that are used to create media products at several companies including the *San Francisco Chronicle*, *Wired* (a high-tech magazine), and several major Web sites. The research was aimed at gaining an understanding of the use of technology *in situ* through interviews with professionals and observations of content creation. They found that producing the media at these companies was fast paced, diverse, complex, as well as routine and production-oriented to meet publication schedules. The work demanded continuous switching between paper and electronic representations and required user mobility to share and distribute drafts. They also observed the following essential functions of paper media: (1) better visibility in terms of resolution and simultaneous viewing; (2) support for rapid informal communication and sketching; (3) service as tokens denoting the transfer of the media from one person to another and designation of the final product; and (4) permanence. The research of the commercial industry offers lessons for the development of hypermedia to support METOC briefings. These lessons demonstrate that paper's essential functions must be reproduced to minimize preparation time and maximize effectiveness critical to METOC facilities such as the Naval Pacific Meteorology and Oceanography Facility at North Island (NPMOF) which is in the forefront of electronically producing and distributing weather information. NPMOF is one of the sites used for this study.

Ideally, the steps in user-centered design should be formalized to produce effective documents. The first and third steps, analysis and testing, are readily formalized. The Pliske et al. (1997) study is an example of a formal approach to the first step and this report is an example of a formal approach to the last step. The middle step, evaluation, however, actually produces the design and might also be formalized, according to Sutcliffe and Faraday (1994). They present a systematic method for the design and analysis of multimedia interfaces and a case study of the development of a shipboard emergency management system. The formal method includes the expansion of the task model to include scripting acts including dialogue statements (Inform, Locate, Emphasize) and information types (Descriptive, Spatial, Operational). The task model they used in their case study was primarily procedural, a limitation discussed in the next paragraph. The media resource requirements are attached to these scripts and organized into a presentation and sequencing structure. Cognitive and attentional guidelines are used to manage the coordination of the media. In the case study, the media produced through this method received favorable reviews from potential users, although there were some usability problems that resulted from the design. Adding a set of usability guidelines would have helped to avoid these problems. Guidelines for general human-computer interaction (HCI) design are available in several sources, and guidelines for hypermedia, including Web page design are being developed (e.g., Borges et al. 1996; Nielsen 1997).

However, the use of guidelines, and the use of a method that is based mostly on a procedural model, will not support the development of an understanding of the present weather situation, the second step in the general forecasting model of Pliske et al. (1997). To model comprehension and situation awareness, an approach such as that of Narayanan and Hegarty (1997) is required. They are developing a comprehension model of how complex machine descriptions are understood using cognitive theory and suggest using the model in the construction of hypermedia manuals. The predictive aspects of their model assist in finding possible sources of user comprehension error. They use a constructive model that proceeds in stages. A critical stage for the present topic is the determination of causal chains of events in the machine's operation, stage three. In the present context, this means that a formal design process that is based on not only a procedural task model but also a weather assessment model must include weather causality chains. This would be particularly important in developing documents to acquaint METOC personnel with the weather in their new assignment area, which was the purpose of the UC-CD. These personnel would not have the mental visual models that experienced, highly skilled personnel have for the area. Fostering the development of these skills would be enhanced by weather-model-based design. With DMARS, the media developers were skilled meteorologists and oceanographers who understood climatological phenomena. Furthermore, weather models were implicitly embedded in some of the dynamic pages through look-up tables computed from weather model output. Thus the document development, as well as the final product, was model-based to some extent.

2.3 METOC Support Documents

NAVOCEANO produces a number of paper products to support the Navy. These documents provide historical meteorological, oceanographic, and geophysical information that impacts Navy operations in a specific geographical area. In general, the documents are tailored to support a specific warfare area such as antisubmarine warfare (e.g., Environmental Guides or Submarine Tactical Oceanographic Reference Manuals), mine countermeasures (Mine Warfare Pilots), or amphibious assault. The documents are image rich. There is descriptive text and the information is organized into chapters, but the chapters consist mostly of graphical images that can be used by METOC personnel to interpret how the current and future METOC conditions may impact operations. The content for these documents is derived from meteorological and oceanographic requirements to support Navy warfare areas and are reviewed by the NAVOCEANO and the supporting warfare commands before being released. A document will have an introduction, a description of daylight duration, winds, tides, rain, currents, temperature, geographical location, sediments, and biology. A more specific warfare section might include acoustic bottom loss, ambient noise, mine burial rates, and acoustic or electromagnetic-system performance predictions. Each METOC unit will have a set of manuals that relate to its regional area.

As mentioned previously, the UC-CD used for this study was the DMARS. Its interface was created following instructional design techniques (Miyamoto et al. 1991) similar to the user-centered design approach above. Instructional design begins with a **needs assessment** that is used to identify the users and evaluate their needs. The users of the UC-CD were identified to be defense planners, platform drivers, antisubmarine warfare (ASW) analysts, and sensor operators. Following the needs assessment is a **task analysis** including an inventory of available data. Next is the **strategy selection** where strategies for accessing the data are developed and a prototype interface is constructed. Finally, the prototype interface is subject to **test and redesign.**

The needs assessment and needs analysis produced the original design goals for the UC-CD. They were stated (Miyamoto et al. 1991) in this way:

> "Design objectives include:
>
> - linking all the parts of the guide into one cohesive whole
> - defining appropriate sequences of information
> - providing branching for quick access to any part of the guide
> - improving retention
> - building in user feedback
> - maintaining flexibility for updates
> - providing self-instruction and "helps"
> - making the program "user-friendly"

The design of the CD-ROM Guide should be **simple, clean, clear, and intuitive**. An intuitive design implies that the user is drawn to the next step with minimal instruction or explanation. The user should not require an extensive instruction manual or computer expertise to run the system. This implies little use of complex commands and maximum use of graphic symbols (icons), multiple windows, and "buttons" that can be "pushed" to go between sections. These new graphic user interfaces should not be difficult to implement with a number of different software programs. Multiple "windows" on a single monitor can each display different information and can be moved and resized enabling data to be compared. The use of "buttons" and "pull-down menus" is essential to maintaining a simple user interface for complex organization. Cross-referencing between window subjects and titles in both a linear index (according to name) and subject index allows for quick access to the information.

The information must be readable on the computer screen. This is not always true for data storage systems where high resolution [300 dots per inch (dpi)] scanned images are viewed on a low resolution (72 dpi) screen. … The text becomes unreadable and many fine lines drop out. It is possible to display only a portion of the image at a time using a "zoom" feature, but this contradicts the speed requirement and makes it difficult to visualize the whole picture…"

These objectives were followed to create a prototype interface that was widely demonstrated and evaluated. The final product, after one redesign, is shown in Fig. 1.

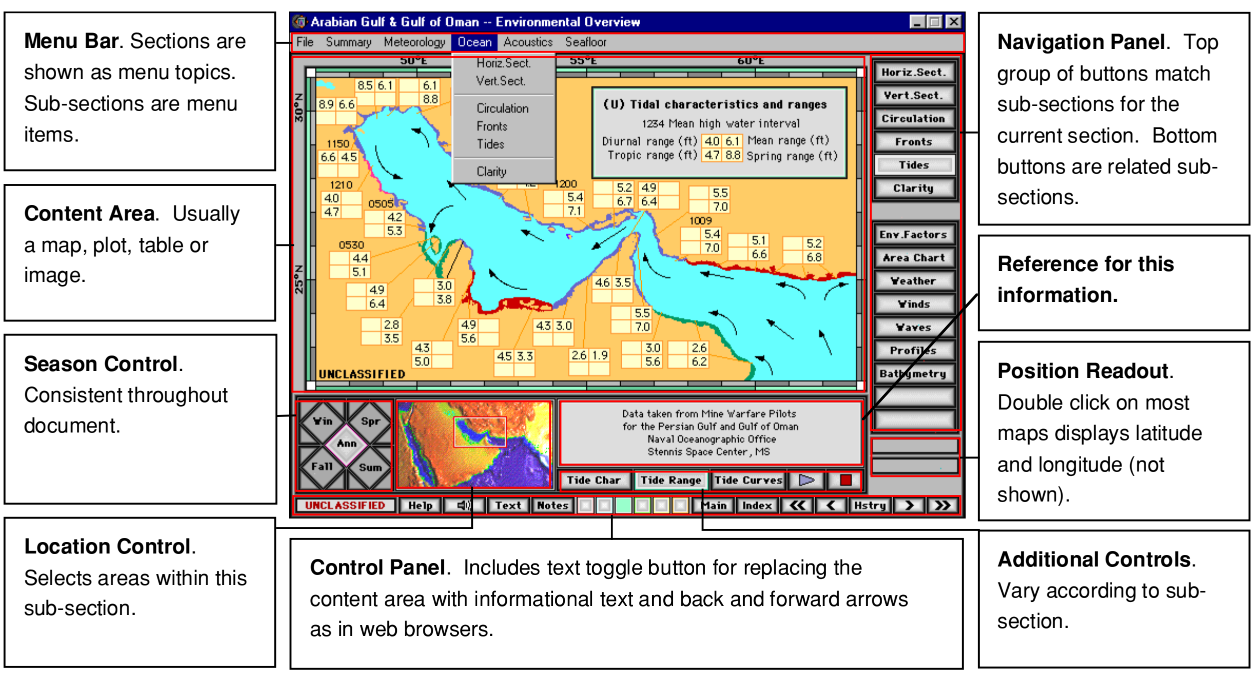

Fig. 1 — A typical DMARS screen with functional descriptions of the main sections. The section borders are highlighted in this illustration.

As put forth in the first design objective, the parts of the guide were seamlessly linked into one cohesive whole through similar appearance and layout. The same base map is used wherever practical. Typically, the only parts of the interface that change from section to section are the content area, the reference for the information, and the additional controls (Fig. 1). The third design objective, providing branching to other sections, was accomplished with the menu bar and navigation panel. Both allow access to other areas with only one user action. The menu bar and navigation panel also provide different access paths into the document, which should speed retrieval, according to the research by Vora et al. (1994).

Making the images screen readable was accomplished by redrawing the graphics specifically for electronic media. The text was rendered with special screen fonts and enlarged. The images were colored for legibility.

3. EVALUATION METHODOLOGY

3.1 Subjects and Equipment

Twelve subjects participated from METOC units at two sites, with six from each site. The two sites had very different missions and computer facilities. Site 1 was an air-support facility with a primary mission of providing meteorological forecasts to pilots. Site 2 was a Fleet-support facility providing meteorological and oceanographic support to air and sea users. Investment in computer facilities was greater at Site 2, particularly in equipment and software for a prototype Web-based system to deliver METOC products to Navy customers. All of the subjects were high-school graduates, of which one had received an associate degree and another had received a master's degree. All of the subjects reported normal vision. The period of both military service and METOC experience ranged from 2 to 15 years amongst the subjects. The average duration of military service in our pool of subjects was just over 9 years, and the average amount of experience performing METOC tasks was 7.5 years. Seven of the 12 subjects reported a degree of geographic familiarity with the Arabian Gulf, the area of interest for this study, having been stationed there. The subjects indicated their experience both with certain software and hardware on a scale of zero to five, zero indicating a complete lack of experience and five indicating current use. The average ratings regarding experience with common computer platforms and software is shown in Fig. 2. Generally, the subjects at Site 2 rated themselves as more experienced with these computer and software technologies, and with the PAPER document used in the study, but the only significant difference was experience with PowerPoint (t (7.5) = -2.45, p = .04; see Appendix A for an explanation of these statistics). None of the subjects had any experience using UC-CD prior to the experiment, although one subject had seen it.

The subjects were familiarized with each document prior to its use. For the PAPER and RP-WEB documents, the subjects read the table of contents. For the RP-WEB, the subjects were also instructed on how to copy images from the RP-WEB document and paste them into brief-preparation software. The familiarity session for UC-CD was the longest because of the lack of prior experience with that particular interface. The session included viewing a tutorial distributed with UC-CD, reading two paragraphs of update notes, instruction on how to obtain a readout of position from maps, and instruction on how to copy images from UC-CD to brief-preparation software. This familiarization lasted about 5 minutes.

The computer equipment at both sites was an IBM-PC-compatible computer running the Microsoft Windows 95 operating system. The computer at Site 1 had 14 MB of memory compared to 31 MB at Site 2. The monitors were estimated to be 17 inches in size and were set to 800 × 600 pixel resolution. At Site 1, the screen was set to 8-bit color mode, and at Site 2, it was set to 16-bit color, although the programs were all designed for 8-bit color and looked the same at both sites. The overall speed of the computer at the second site was observed to be faster than that at the first site. This created improved performance for both the RP-WEB and UC-CD documents at Site 2.

The response time of the network used for the Web connections was estimated to be equivalent at both sites and did not cause observable delays in Web download time when the network was up. However, the network did go down two times at Site 1 and seven times at Site 2. Most of the delays were fairly short, ranging from one minute to half an hour. However, there were also three long delays of several hours to overnight. The long delays twice impacted the testing schedule, in one case requiring that an additional trip be made to the test site. All network down time was excluded from the analysis of problem completion time.

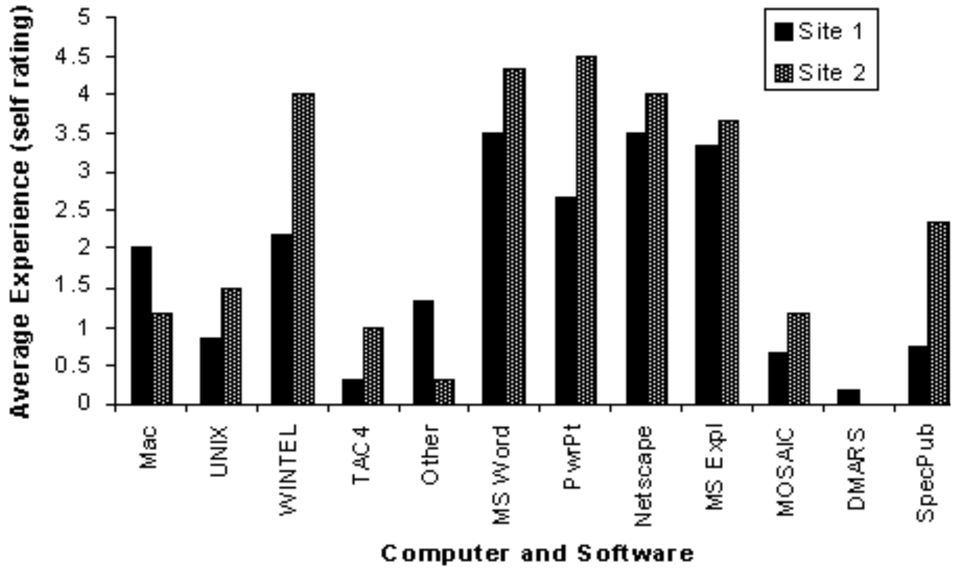

Fig. 2 — Self-reported software and computer hardware experience at the two sites

3.2 METOC Documents Evaluated

Three fielded documents covering the Arabian Gulf were used in this research: 1) a paper document (PAPER); 2) a Web version of the paper document (RP-WEB) with a slightly modified geographical coverage area; and 3) the DMARS UC-CD described in the previous section. All three documents are NAVOCEANO products. We chose these documents because the information is similar in each of the documents and is derived from the same sources. UC-CD has the "look and feel" of a multimedia CDROM, and the RP-WEB used in our testing is typical of Web documents. All three documents include an environmental summary for an area and details of the meteorology, oceanography, bottom characteristics, and ASW information. Table 1 lists typical topics. Each document is done in a style that typifies the interfaces found in each media.

Although the information contained in the documents is essentially the same, the documents differ in navigation through the information and in some cases they differ in how information is accessed from a particular interface, once the information has been found.

3.2.1 Navigation Methods

The primary method of navigating all three METOC documents is by topic. For PAPER and RP-WEB documents, the topics are listed in the table of contents, which is the first page in each document. In UC-CD, the main topics are menu headings on the pull-down menu bar and subtopics are items in each menu. Thus, with one action (selecting an item from a menu) users can go directly to the subtopic of interest. Furthermore, the menu bar is accessible anywhere in the document, whereas with the PAPER document and the RP-WEB, users must first go to the table of contents page before they can find a new topic. In the PAPER document, the table of contents gives page numbers to turn to for each topic. In the RP-WEB document, the topics are hyperlinks that bring up the topic directly. In addition to a table of contents, the RP-WEB document also has a list of figures with links directly to the figures and a list of tables with links directly to the tables.

Table 1 — Typical Contents of a METOC Document

• **Environmental Summary**	• **Oceanography**
— Location	— Sea Surface Temperature
— Critical Factors	— Currents
— Location	— Surface Currents
— Weather	— Ocean Fronts and Eddies
— Winds	— Vertical Sections of Temperature, Salinity, and Sound Speed
— Waves	— Biology
— Ice Conditions	— Bioluminescence
— Circulation	— Marine Mammals
— Tides	— Biological Sound
— Fronts and Eddies	— Attenuation of Light
— Bathymetry	• **Bottom Characteristics and Magnetics**
— Physiography	— Physiographic Provinces
— Acoustic Conditions	— Bottom Sediments
— Propagation Loss	— Sediment Types
• **Meteorology**	— Bottom Loss
— Climatology	— Magnetic Anomaly Detection (MAD)
— Duration of Daylight	• **Acoustics**
— Location of Coastal Meteorological Stations	— Sound-Speed Provinces and Profiles
— Seasonal Air Temperature	— Range to Convergence Zone
— Precipitation	— Topographic Noise Stripping
— Cloud Cover	— Ambient Noise
— Visibility	— Volume Reverberation
— Surface Winds	— Propagation Loss
— Storm Tracks	• **Air ASW Acoustic Conditions**
— Wave Height	• **Surface ASW Acoustic Conditions**
— Atmospheric Ducts	

There are some additional methods of navigation besides by topic. The PAPER document can be browsed by simply flipping through the pages. The UC-CD has a navigation panel on the right side of the screen that lists context-sensitive links to related subtopics. These links are updated as the user browses the UC-CD and take the user directly to related information, bypassing the topical menus. In addition, both the RP-WEB document and the UC-CD have a back button to return to the previously viewed page. These lists were also included in the PAPER document. None of the documents had an index or full document search capability. The closest to these were a topical overview page in the UC-CD, and in the RP-WEB browser, a capability for word search within the loaded page.

3.2.2 Interfaces to Information

Some information was presented in the same format in all three documents. For example, the tide range is keyed as four numbers in a grid for several locations on a map. Mean range is in the upper right quadrant, followed clockwise by the spring, tropical, and diurnal range. Although the information is essentially identical, there are still differences inherent in the media itself. One important media difference is that paper has much higher resolution than most computer monitors. The design of the UC-CD compensated for this reduced resolution in the tidal grids by using a larger font for legibility. As shown in Fig. 3, tides for the entire gulf were shown on a map that only used a portion of a 640 × 480 monitor. If the scanned image from the PAPER were shrunk to the same size in the computer documents, the fonts would be illegible on a normal computer monitor.

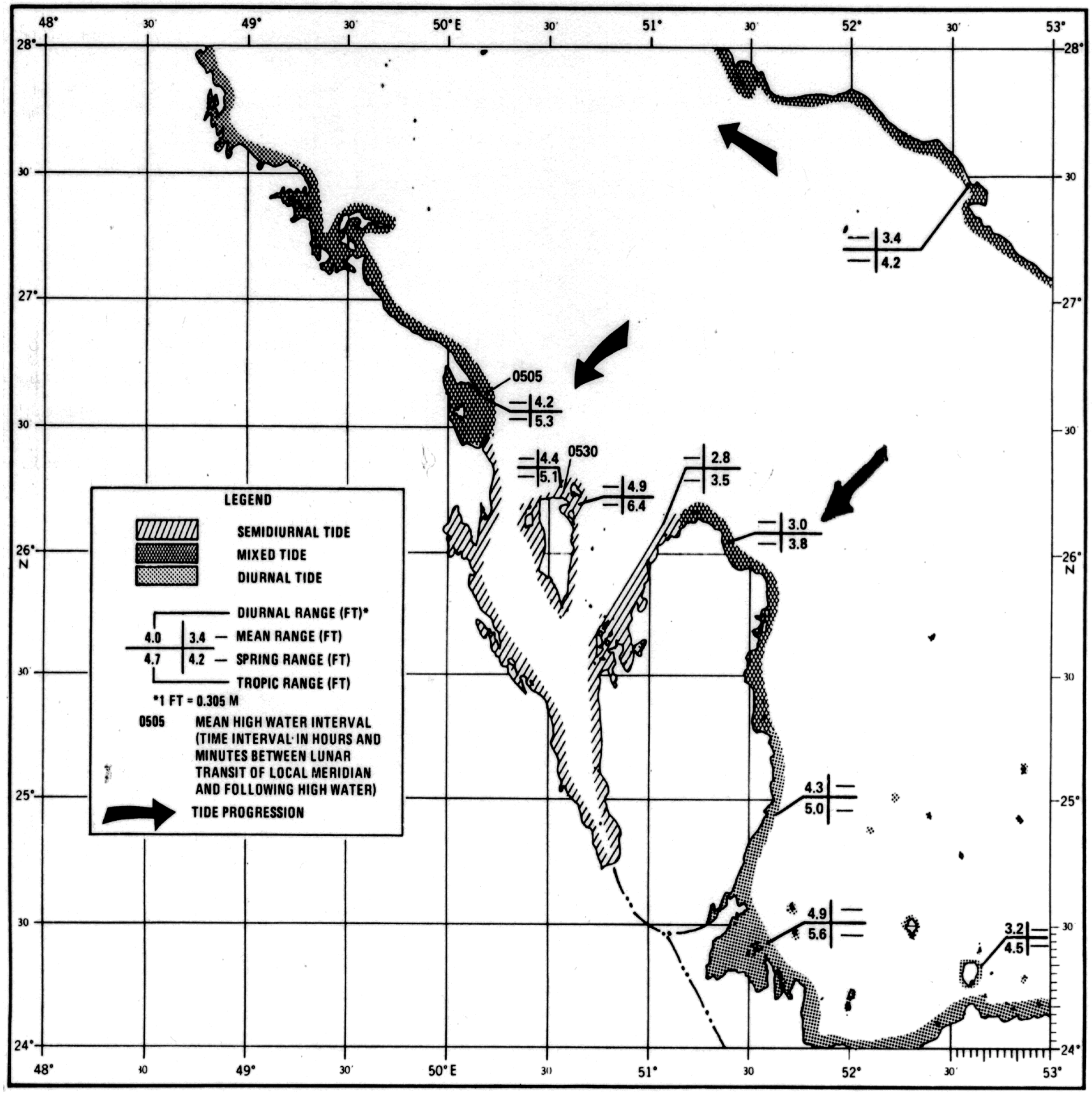

Fig. 3 — Illustration of how tidal range data were presented in the PAPER and RP-WEB documents

Other information was presented in a different form in the three documents. The winds and sound speed profile interfaces, for example, are quite different in the UC-CD, compared to the PAPER and the RP-WEB. Furthermore, the RP-WEB and PAPER documents used on Problems 1 through 4 of this study did not have information on daylight duration. So for Problem 5, we used the daylight chart published in another METOC document and available at another METOC Web site. Since this problem did not involve browsing (i.e., the subject began the problem by viewing the daylight page in all three documents), this change did not confound the design or analysis. The main reason for switching to other documents was to test the redesigned interface on the UC-CD because it is so different (i.e., compare Figs. 4 and 5). In the PAPER and RP-WEB documents, duration of daylight is found by locating the date of interest on the horizontal axis and the latitude of the place of interest on the vertical axis, then reading the duration for the curve that is at the intersection of the date and latitude. In the UC-CD, the daylight duration is computed whenever the user selects a location on the map. However, the interface presents the user with a default duration when the page is loaded, which can and did produce errors if the user does not explicitly click on the correct location.

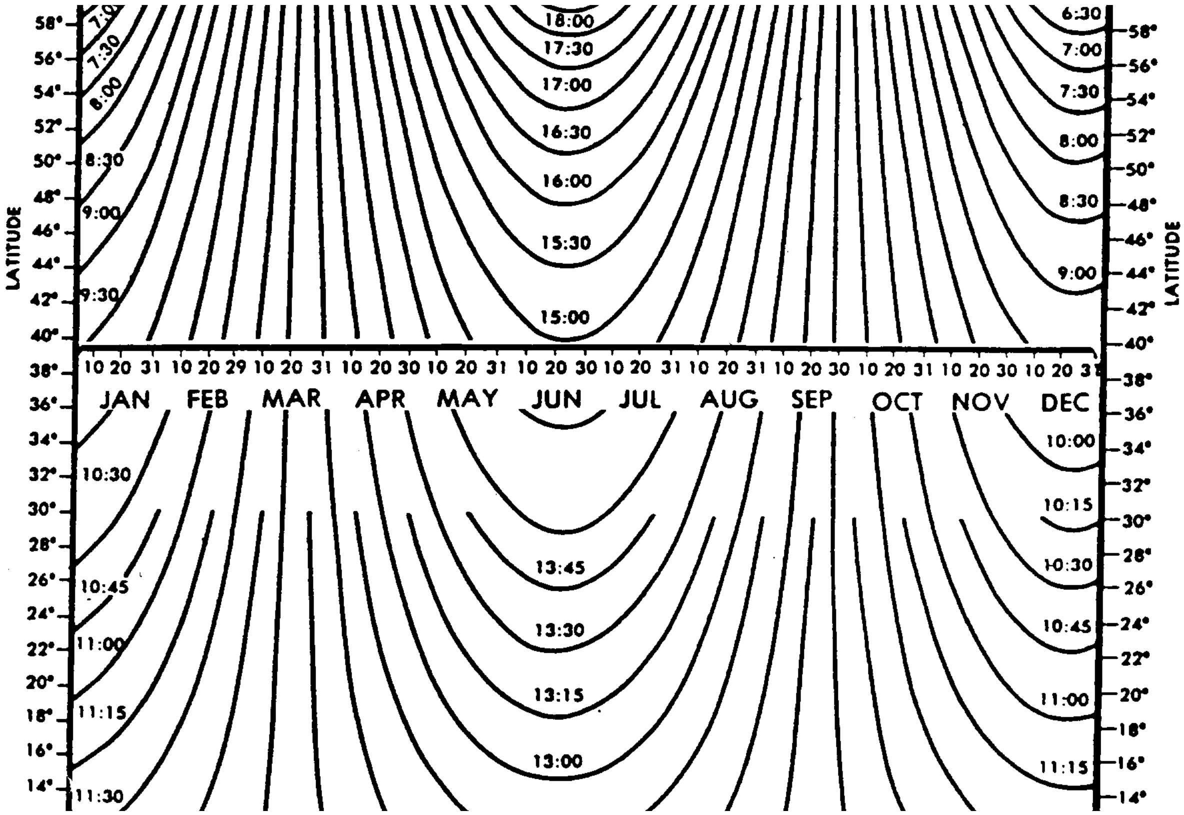

Fig. 4 — Illustration of how daylight data were presented in the PAPER and RP-WEB documents

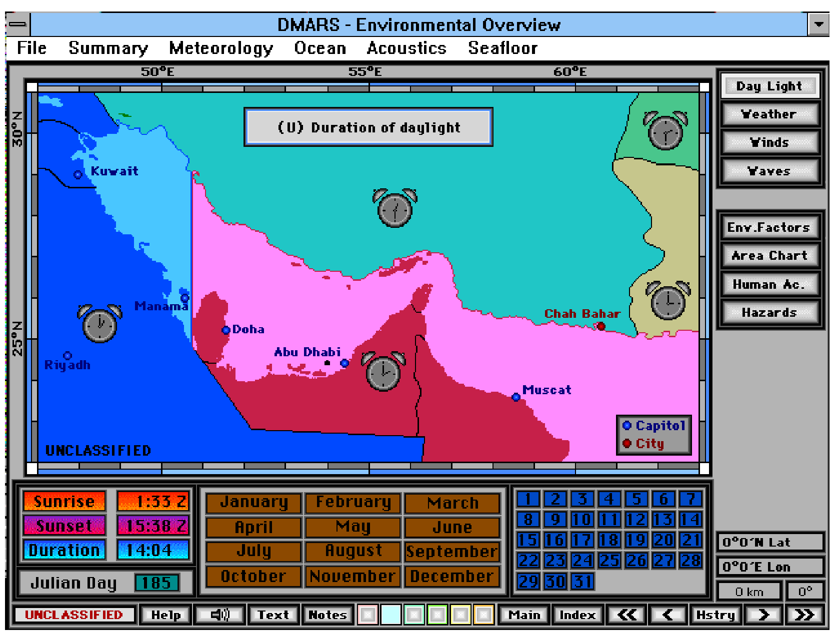

Fig. 5 — Daylight interfaces from the UC-CD

3.3 Problems

The subjects performed five problems that are common in METOC operations. These tasks were designed to represent three types of METOC activities: (1) preparing weather briefing, (2) searching for information, and (3) analyzing information. These problems were selected based on observations of METOC personnel during a Fleet exercise and through conversations with METOC personnel. Pilot testing was used to refine the procedures, especially for Problem 1.

The subjects worked freely through each of the problems. We were able to record the time spent on various tasks using a real-time activity recorder called Activity Catalog Tool (A.C.T.), a product that runs on Macintosh computers (Segal and Andre 1993). The types of tasks we logged are described in detail in Section 3.5. Observations recorded with the A.C.T. tool were used to categorize the total time spent on specific tasks including Browse, Interpret, Prepare Brief, and several other tasks. Table 2 and the discussion following summarize the measurement objectives for each of these tasks.

One of the primary duties of Navy METOC personnel is to prepare weather briefings, and Problem 1 was designed to address this work requirement. The subjects were asked to prepare a weather briefing for a particular location including an image showing the site location and regional bathymetry; tidal information including characteristic tide, tide range, and typical tide curve; and wind information including prevailing direction and speed for spring and autumn. It was designed to measure the overall time to search the METOC documents for the information and to assemble the information into a brief using a word processor or briefing preparation software. This problem was the least structured of the five, but had been developed through pilot testing to ensure that it was feasible. During this problem, we were able to observe the subjects in a setting similar to where they would normally work. The subjects were instructed to include images as appropriate and explicitly asked to include an image of the bathymetry. For electronic media, these images could be copied from the METOC document and pasted into the briefing software. For the PAPER, the subjects were asked to put placeholders in the briefing document or the PAPER document and, after the session, were asked how they would move the paper images into the brief and timed doing their preferred procedure for one image. This measured time for one image was then multiplied by the number of images that had been tagged for inclusion in the brief in order to produce a total time to generate images for the briefing. Pilot testing indicated that this problem could be completed in about 30 minutes. By cataloging activity, we were able to separate navigation time, interpretation time, and brief preparation time.

The rest of the problems were designed to focus on selected aspects of METOC information retrieval, relying on the following assumptions. Accessing information involves searching for the information, navigating to it, and then reading and interpreting what is there. Multimedia may assist in both parts. If the location of the information is known, then accessing it will include the time needed to navigate to the information, and read it, but not time to find it. When the location is not known, accessing it will also include time to search for the information. If the information is presented in a relatively simple form, or a form that is similar across the documents, then interpretation time should be minimal and not vary. Finally, search and navigate time can be eliminated by starting the subjects at the location of the information in the document.

Problem 2 asked the subject to retrieve the same tidal information as in Problem 1 for a different geographic location. Having completed Problem 1, the subject already knew where this information was located. This problem allowed us to measure navigation differences by assuming the search time was negligible. The actual statement of the problem was: "Record the mean tide range near (a location)." However, if the subjects failed to find the specific tidal information in Problem 1, and realized this, then Problem 2 could include time to find the information in a second search attempt.

Table 2 — Measurement Objectives for the Tasks Observed on Each Problem

<table>
<tr><th rowspan="2">Problem</th><th colspan="2">Access Information</th><th rowspan="2">Prepare Brief</th></tr>
<tr><th>Browse</th><th>Interpretation</th></tr>
<tr><td>1. Prepare weather brief including bathymetry, tides, and winds</td><td>Search and navigate differences</td><td>Interpretation differences (includes media and interface differences)</td><td>Brief preparation differences</td></tr>
<tr><td>2. Answer question about tides</td><td>Navigate differences</td><td rowspan="2">Pure media differences (three documents had similar appearance)</td><td rowspan="4">Not applicable</td></tr>
<tr><td>3. Answer question about fishing</td><td>Search differences</td></tr>
<tr><td>4. Compare sound speed profiles from different seasons</td><td rowspan="2">Not applicable</td><td>Sound speed profile interface design differences</td></tr>
<tr><td>5. Determine duration of daylight</td><td>Daylight interface design differences</td></tr>
</table>

Problem 3 asked for fishing activity, something not previously requested and therefore requiring a new search. This problem was designed to measure the search component of browsing time, but by necessity also includes the interpretation time. The actual statement of the problem was: “Is there any shrimp trawling at (a location)?”

In Problems 2 and 3, the design of the interface is generally the same for the different documents (a coded map). But again, there are differences in the presentation of the information to the user. The PAPER version is a black-and-white book. The RP-WEB version was scanned from the PAPER version, but when the scanned images are displayed on the lower resolution computer monitor, the images required enlargement to be legible. Therefore, the image and the accompanying text could not be simultaneously displayed. The UC-CD also started with scanned images, but they were redesigned; color was added to the images and the writings on the images were made larger in computer fonts so that the overall size of the image did not have to be increased as much. These presentation effects could be measured from Problems 2 and 3 by looking at the interpretation time, which was recorded separately from the search time using A.C.T.

Problem 4 asked the subject to compute the difference in sound speed for summer vs winter and Problem 5 asked the amount of daylight. Unlike Problems 2 and 3, the design of the interface significantly differed between the UC-CD and the other documents for these problems. These problems were designed to measure the effect of the design differences. In both of these problems, the subject started at the screen that had the information on it, so the only measure was the interpretation time. The statement of Problem 4 was: “How much faster is the sound speed in summer than in winter at a depth of 10 meters at (a particular location)?” The statement of Problem 5 was: “What is the duration of daylight at (a particular location)?”

A time limit was not used for any of the problems. Total participation by an individual subject lasted no more than 3 hours. A questionnaire was administered to gather background information about the subject’s METOC experience and a debriefing was given at the completion of the data collection.

3.4 Statistical Design

Each subject completed these problems using each of the three documents described earlier. This within-subjects design was required because of the low number of military METOC personnel that would be available at each site. In most cases, evaluation of each subject was completed in one day and the subject finished all three of the documents before starting the next subject. However, due to network problems mentioned previously, in one case a subject was started before the previous one used his third document. In another case, the testing had to be aborted and restarted the following week. In order to reduce the statistical effects of this design, a variation of a Latin squares design, called digram-balanced, was used. In this design, the order of testing with the three documents is controlled so that each document is preceded and followed by every other document one time in the subject population, at each of the two sites. For three documents, this required six testing orders, establishing the need to test six subjects at each site. When switching from one document to another, the details of the problems were changed.

3.5 Performance Measures

Both process and outcome measures were obtained. Process measures included the time taken to complete the problem, including the time spent on specific tasks that were required, such as browsing the document. These specific tasks are described in the following section. Process measures also included how often specific tasks were initiated for each problem and the average time spent in a specific task. Outcome measures included the number of images used in the briefing for Problem 1 and the accuracy of the answers in Problems 1 through 5.

3.5.1 Process Measures

As described previously, completion of the problems required specific tasks. The time spent in the tasks were recorded in real time using A.C.T. It uses up to nine keys in the home row to represent separate tasks the user can perform. Pressing a key records the start time of the task. Pressing it again records the end time of the task. If the tasks are defined to be mutually exclusive, then pressing any task key records the end time of any ongoing task, and the start time of the task represented by the key. In this experiment, four tasks were recorded representing subcomponents of two general tasks: access information from the document, and assemble a document (in this case a briefing) that summarizes the information. The specific tasks were defined as follows:

Access Information:

Browse: search for topic in the METOC document by navigating from one section to another to another;
Interpret: read information from a specific section.

Prepare Brief:

Compose/edit: generate and/or modify the brief using a word processor or presentation software;
Copy/paste: copy material from the METOC document into the brief.

In addition to these tasks, four other keys were used to record the following:

Instruct: read the task instructions;
Offtask: engage in an activity not related to the experiment;
Wait: wait for Web page download;
Other event: record any other events or activities (when this key was pressed, the observer had the option of typing a short note to describe the event).

The difference between Browse and Interpret depended on whether the subject had navigated to the correct topical section of the document. For example, in Problem 2, Browse would terminate when the subject reached the section on tidal information. Interpret would then begin. Reaching the tidal section does not mean the subject had reached the actual page or screen where the answer (e.g., tide range) was located. The copy/paste operations were performed differently depending on the document. For the UC-CD, the users used a Windows utility that can capture a snapshot of the screen into a temporary buffer. This image can then be pasted into the document the user is producing. The keystroke operations for this task are simple—using a three-button mouse, the user presses the third button and wipes the section of the image to be captured. Upon lifting the mouse button, the image is captured. After switching to the document being edited, the user can paste this image by selecting paste from the edit menu or by pressing two keys simultaneously. When using the Web browser, upon pressing the right button, the browser produces a pop-up menu with an option to copy the image. This can be pasted into the document preparation software.

Table 3 lists the specific tasks that were relevant for each of the five problems. As explained in the methodology section, the five problems were designed to elicit different combinations of the tasks. All problems required interpreting the METOC material. The first three required browsing for the material, although for Problem 2, this search was to a location that had already been visited. On Problems 4 and 5, the user began the problem at the page with the salient information. Because only Problem 1 required generating a document (i.e., a weather briefing), it alone required the subjects to engage in the tasks of *compose/edit* and *copy/paste*.

Table 3 — A.C.T. Tasks Recorded for Each Problem, Grouped by the Measurement Objectives Listed in Table 2

	Task						
	Access Information		**Prepare Brief**		**Other[1]**		
Problem	**Browse**	**Interpret**	**Compose/ Edit**	**Copy/ Paste**	**Instruct**	**Offtask**	**Event**
1	X	X	X	X	X	X	X
2	X	X			X	X	X
3	X	X			X	X	X
4		X			X	X	X
5		X			X	X	X

[1] These tasks were included to record supplementary information about the subjects' behavior

For each task, two dependent variables were derived from the A.C.T. records: task time and task frequency. Task time refers to the total time spent in the particular task for a problem. Task frequency refers to the number of times the particular task was initiated during the problem. One additional measure, total problem time, was computed by adding the task times for a particular problem, excluding *Offtask* and *Instruct* time. Delays caused by network disruptions were not included in any of the time measurements. For these three measures (total problem time, task times, and task frequency), an analysis of variance (ANOVA) was used to determine whether the independent variables of document, site, or order had a significant effect on performance. A General Linear Model was used with the variables and degrees of freedom listed in Table 4. Appendix A provides details on the statistical tests. Subjects were defined as a nested variable within the site variable.

3.5.2 Outcome Measures

Outcomes from the five tasks were measured to determine how well the tasks were performed. The brief prepared as part of Problem 1 was graded for accuracy and content before starting Problem 2. As part of this evaluation, the subjects were requested to present the brief in an abbreviated form to the

Table 4 — ANOVA Design

Source	Degrees of Freedom
Document type	2
Site	1
Order	2
Subject (nested within Site)	10
Document x Site	2
Site x Order	2
Error	16

experimenter who used a scorecard (see Appendix B) to grade the briefs and record the number of images used. This generally took less than 5 minutes. After grading Problem 1, the subjects worked the four fill-in-the-blank problems (Problems 2 through 5). The experimenter recorded the answers to these problems on the scorecard and graded them at the end.

The brief was graded pass/fail on eight items and the four fill-in-the-blank problems were each graded pass/fail on one, for a total of 12 points. Information that was missing but had been requested in the problem statement was graded incorrect. In the brief, one point was earned for an image of bathymetry that included the location requested in the problem statement. Three points were earned for tidal information, one point each for the characteristic tide (diurnal, semi-diurnal, or mixed), tide range (a number) and typical tide curve (an image). Four points were earned for wind information, one point each for wind direction in the spring, wind speed in the spring, wind direction in the fall, and wind speed in the fall.

Images alone were insufficient to receive credit for the characteristic tide and tide range. If the subject did not volunteer the exact information for the location specified in the problem statement, then the experimenter requested the information from the subject. If the answer could not be provided or was incorrect (e.g., for the wrong location), it was graded incorrect even if the correct information was contained on the image provided. In contrast, for the four wind points, credit was given if an image or table was included that contained the information, regardless of whether the correct answer was read from the table or image. The leniency was granted to prevent a bias toward the UC-CD. The UC-CD wind interface, consisting of an arrow indicating wind direction and a number indicating speed, left little room for error. But the PAPER and RP-WEB interface for winds was a table with percentage winds for each direction and speed with considerably more numbers. Grading the fill-in-the-blank problems (Problems 2 through 5) was straightforward. Problems 2, 4 and 5 had numeric answers that had to be close to the correct answer to receive credit. Problem 3 was a yes-no answer.

4. RESULTS

4.1 Effects of Document Type on Task Time

4.1.1 Browse Time

Our primary interest is in the effects of document type on performance of the different tasks. One of the consistent results across several problems is that document type affected Browse time on the three problems—1, 2, and 3—that required searching for information. On each of these problems, document type had a significant effect on Browse time. These results are shown in Figs. 6 through 8,[1] and the

[1] Each figure has standard error bars that illustrate the variance of the average.

statistical tests are summarized in Table A5. It might be noted that all of the significant differences involve comparisons with the PAPER, the traditional form of METOC information.

One of the consistent outcomes shown in Figs. 6 through 8 is that the RP-WEB was slower than the PAPER. On the problem that had the longest overall Browse time, Problem 1, creating the briefing, Browse time was significantly slower with the RP-WEB, compared to either the UC-CD or PAPER (Fig. 6). This result is important because this problem was representative of the types of briefings that the METOC community does on a regular basis. This problem also was least structured and the subjects were free to locate several types of information in any order they chose.

As shown in Fig. 7, for Problem 2, answering tidal questions, the RP-WEB was slower than PAPER, but the result is marginally significant with a probability of 0.10 (Table A5). However, the pattern of results is similar to the pattern for Problem 1. The recurrence of this outcome was surprising because the subjects had already retrieved tidal information for another location on the first problem. Despite their recent experience in locating tidal information, the slower performance on the RP-WEB persisted.

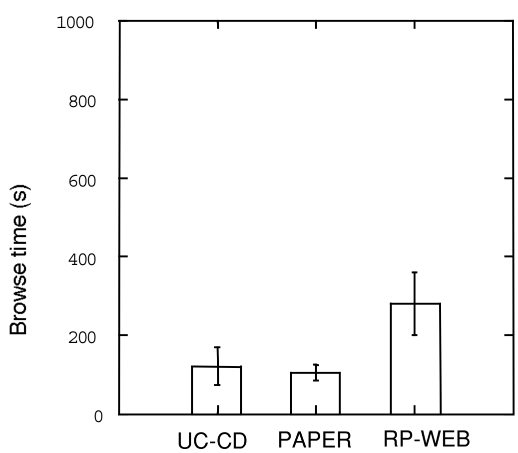

Fig. 6 — Problem 1: average Browse time by document type

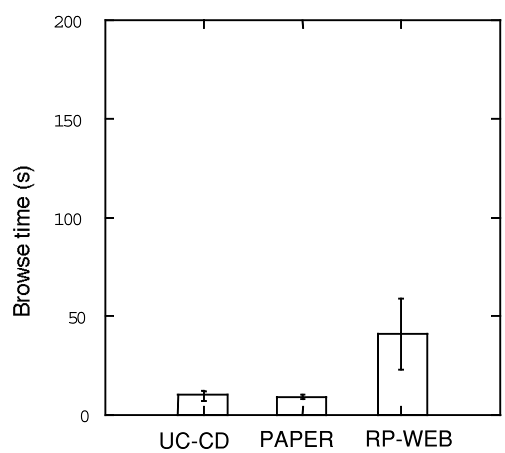

Fig. 7 — Problem 2: average Browse time by document type

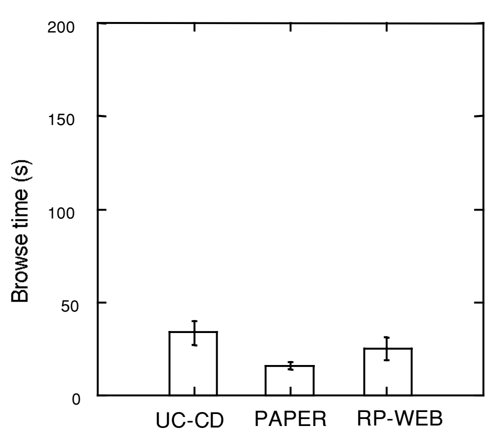

Fig. 8 — Problem 3: average Browse time by document type

On Problem 3, the subjects had to retrieve information about fishing activity, and although the RP-WEB browse time was slower than PAPER, the significant result here is the slower Browse time using the UC-CD, compared to PAPER (Fig. 8). This result was probably due to the UC-CD menu design that placed "Fishing" as an element on the menu entitled "Human Ac." In the other documents, fishing activity was probably easier to locate because it was explicitly listed as a separate topic in the table of contents. This was the first time that the subjects had been required to find information about fishing activity. Indeed, this was the first time that the subjects had to go to the "Human Ac." menu in the UC-CD. The longer times for the RP-WEB[2] on these problems cannot be attributed to the Web page update since network delays were negligible and the experiment was suspended whenever the network was down. Furthermore, the subjects started at the home page for the METOC information and simply had to browse within the RP-WEB document.

4.1.2 Total Task Time

Document type also affected the total time to complete Problem 2. The total time on this problem was significantly longer for the RP-WEB than for the PAPER (Fig. 9). As noted previously, part of this would have been due to the Browse task, but the addition of the other task times increased the effect. Interpret was the other key primary task for Problem 2 (see Table 3 for the listing of tasks by problem). It is apparent that browsing for and interpreting tidal information took significantly longer with the RP-WEB compared to PAPER.

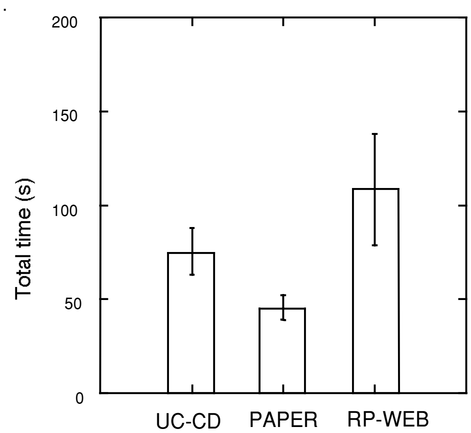

Fig. 9 — Average total time on Problem 2
by document type

The probable cause of the longer RP-WEB time for Problem 2 is suggested by another result, a significant document-by-site interaction on total time, shown in Fig. 10. The important result in this interaction is that the total time on Problem 2 using the RP-WEB at Site 1 was significantly longer (see Table A3) than the other documents at Site 1 and significantly longer than every document at Site 2. As described previously, there were functional differences between Sites 1 and 2. The METOC facility at Site 1 was an aviation support detachment, whereas the facility at Site 2 supported both aviation and surface customers. It is therefore likely that the personnel at Site 1 were less familiar with tidal information, and when given this task, did not know exactly where to find it. Thus it is understandable why they might take longer to find the tidal range information. But the longer time occurred primarily

---

[2] As seen in Fig. 8, Browse time on Problem 3 was slower for the RP-WEB compared to the PAPER, but this difference was not statistically significant in the post hoc test (Table A5). However, it is consistent with the results of Problems 1 and 2.

when they used the RP-WEB. This suggests that using the Web can produce significant delays in information retrieval when a person does not know specifically where to find the information, given the design of the Web document in this instance. Interestingly, the difference between the RP-WEB and PAPER interface was relatively minor for tidal information. For both, the tidal range information was in a map with small, boxed tables in a 2 by 2 format (see Fig. 3). This table was inserted into the PAPER document close to the text that refers to it. The figure was also included in the list of figures. In the RP-WEB document, the figure was nearly identical, and is hyperlinked to the text that refers to it. It is also included in the list of figures. This result suggests that simply transferring information from paper to Web may increase the time it takes to retrieve information.

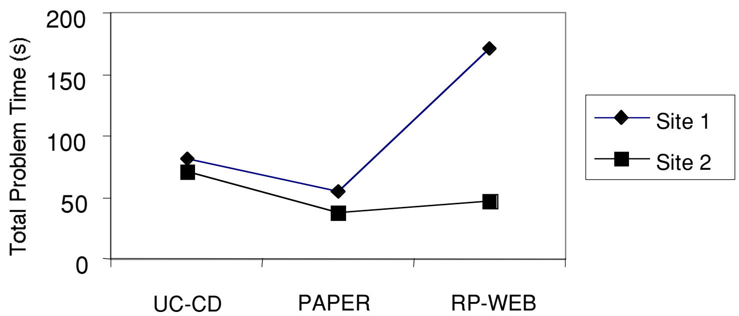

Fig. 10 — Interaction of site and document type on average total time on Problem 2

Total time on the other problems did not vary significantly by document. For Problem 1, brief preparation, we found that completing the total task was significantly quicker with the PAPER before we adjusted for the time needed to actually prepare the transparencies. Once we added the time estimated to complete the physical transparencies, the three documents were similar in total task time. This result is important because it may be easy to overlook the incidental tasks needed to complete a briefing using paper documents.

4.1.3 Interpret Time

There was one significant effect of document on Interpret time. This occurred with Problem 5, determining the duration of daylight at a particular location for a particular time, and was not surprising given the differences in the documents. The effect is shown in Fig. 11; using the UC-CD was significantly faster than using PAPER. This was the only result where either UC-CD or RP-WEB was faster than the traditional document, PAPER. On this problem, Interpret was the only relevant task because the users started at the appropriate page in the document, and did not produce a document. This meant that the subjects had to analyze the information presented for daylight duration.

4.2 Effect of Document Type on Number of Briefing Images Used in Problem 1

One additional measure was analyzed using the ANOVA design in Table A4. This was the number of images that the subjects decided to include in their briefing. The results indicated an effect for document at the 0.11 level (F (2,16) = 2.56, p = .11, see Appendix A for an explanation of this test). Post hoc tests did not support any significant comparisons at this level, although as shown in Fig. 12, the most images were generated when the UC-CD was used.

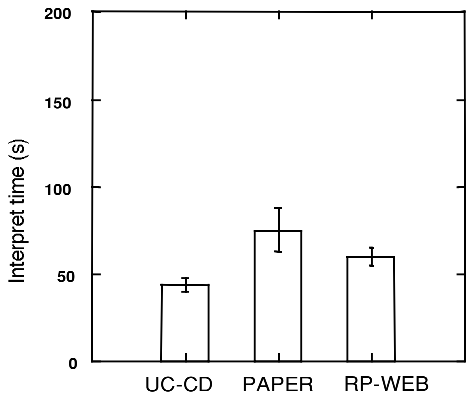

Fig. 11 — Average Interpret time on Problem 5 by document type

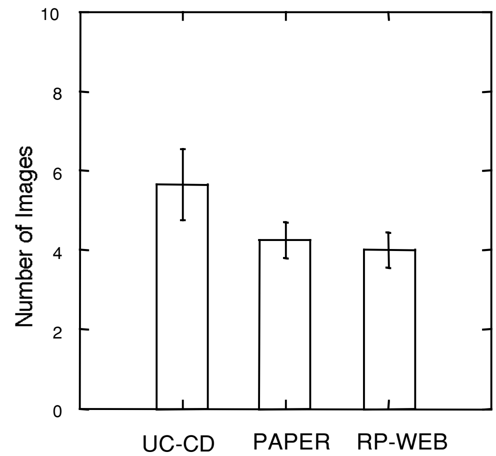

Fig. 12 — Number of images used in the briefing prepared in Problem 1 by document type

4.3 Effect of Document Type on Accuracy

Overall accuracy on the problems was only about 80%. This was unexpected and was probably due to several factors. Accuracy was lowest with UC-CD as shown in Fig. 13. Summing across all subjects, with UC-CD, there were 41 errors in 144 answers; there were 23 errors each on the RP-WEB and the PAPER. The higher error rate with UC-CD is not surprising given that it was a new document for all the subjects, and that the training with UC-CD was minimal. An ANOVA with total errors on all problems as the dependent variable showed significant effects for document (Table A5). The ANOVA also produced a significant effect for order with fewer errors on the second document compared to the first (Table A5), which is clearly a learning effect.

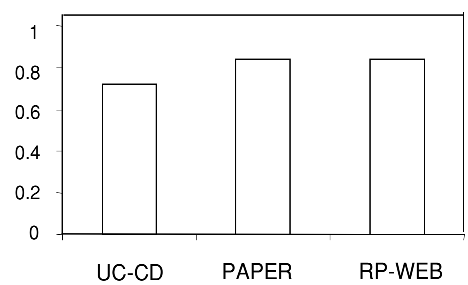

Fig. 13 — Average accuracy on all problems by document type

The error rate varied by problem, but the data were insufficient to do an ANOVA for each problem, so only the actual error counts are described. For the bathymetry part of Problem 1, fewer errors were made with UC-CD than either RP-WEB or PAPER. For the tides part of Problem 1, there were a few more errors with UC-CD (15, 12, and 13 for UC-CD, RP-WEB, and PAPER respectively). For the winds part of Problem 1, there were far more errors with the UC-CD (10 errors compared to 0 for the RP-WEB and PAPER). There were more errors on the UC-CD on Problem 4 (sound speed), with 8 of the 12 subjects making errors compared to 3 each for RP-WEB and PAPER. There were also more errors with

UC-CD on Problem 5 (daylight duration), with 4 of the subjects making errors compared to 0 for RP-WEB and 2 for PAPER.

4.4 Miscellaneous Results: Testing Order and Site

As expected, the order in which the three documents were used had a significant effect on aspects of task completion time on every problem, and had an effect on task frequency for the first problem (Table A4). The subjects were generally slower in using the first document on all problems and did more task switching than on the other problems. Because of the digram-balanced design, each document was in the first position equally often, and each document was followed by the two other types of documents equally often. It might be useful to know whether there is an interaction of order and document type. However, one of the limitations of a Latin square design is that this information cannot be assessed.

There was a significant effect of site on Problem 3 (fishing activity): search time by the subjects at Site 1 on this problem was significantly longer than search by subjects at Site 2.

5. DISCUSSION

5.1 Conclusions

Several conclusions are drawn from the results.

1) The RP-WEB document is slowest to browse.

Problem 1 is the best indicator of overall browse time because it included a mix of activities a METOC officer would do to prepare a brief. In Problem 1, the RP-WEB document was slower to browse than either the PAPER document or the UC-CD. In Problem 2, retrieving tidal information from a known location in the document, the RP-WEB document was slower in total time and showed a tendency to be slower than the PAPER document ($p = .07$). It was also slower, on average, than the UC-CD document although not significantly so. The RP-WEB took 41.4 s, whereas the UC-CD only took 9.8 s, compared with PAPER at 9.3 s.

The only problem that tested browse time where the RP-WEB document was not significantly slower than the PAPER document was Problem 3. In Problem 3, the UC-CD was significantly slower than the PAPER document. This may have been because of the wording used in the UC-CD. Problem 3 asked for shrimp trawling at a particular location. In the PAPER and RP-WEB documents, this information was located in a section called “Fisheries and Fishing Craft” whereas in the UC-CD this information was located in a section called “Human Activities.” Fishing is only one category of Human Activities, thus fishing is more specific to shrimp trawling than is Human Activities.

The main advantage to browsing the PAPER document over the RP-WEB document seemed to be the ability to quickly scan the information in the entire document by flipping the pages and glancing at images. The sequential order in which the information is presented allowed the user to remember where one piece of information was relative to another, thus allowing him to jump from one section to another without always starting at the beginning. This advantage can be overcome with a computer interface as is evidenced by the UC-CD interface. It was not significantly slower than the PAPER document on the first two problems. The UC-CD simulated the sequential nature of the PAPER document with the menu bar at the top, as illustrated in Fig. 1. The sections are in order from left to right. The user can “flip through” the sections and see the subsection names by clicking down on the menu bar and sliding the pointer across to read all the section and subsection names. In addition, the navigation bar on the right side of the UC-CD screen contained links allowing the user to jump to related sections similar to paging forward or backward through a paper document. It was observed that some of the subjects performed almost all of their navigation in this manner with very minimal use of the menu bar.

2) *The daylight interface in the UC-CD document was faster to use than the one in the PAPER document.*

The daylight interface in the PAPER and RP-WEB documents is a graph where the duration of daylight is determined by picking one curve from a family of curves for the particular day of year and geographic latitude. Several subjects appeared puzzled when they first looked at the graph, but most were able to quickly figure it out. Therefore, it did not seem to be very intuitive. Once understood, it is a three step process: 1) find the day of year on the horizontal axis, 2) find the latitude on the vertical axis, and 3) determine which curve is at the intersection of the day of year and latitude. The label on the curve is the duration of daylight. In contrast, the subjects immediately grasped how to operate the daylight interface in the UC-CD. The steps are to select a geographic location by clicking on a map. Then select the month from a list of months and the day from a list of days and then read out the duration of daylight.

Although the subjects were faster using the UC-CD than the PAPER document to produce a brief, more subjects made errors using the UC-CD than they did using the PAPER document (67% correct for the UC-CD, 83% for the PAPER document). The primary cause of error was failure to set some of the controls (location or day of year). This may be alleviated by displaying the location and day of year with the duration of daylight readout.

3) *There is a tendency toward more errors with the UC-CD than with the PAPER or RP-WEB documents. There is also a tendency toward fewer errors with the second document compared with the first (at alpha = 0.1).*

Experimenter notes suggest that many of the errors came from not selecting the location on the map. The interface was not designed to prompt or mandate this user input, as should be done to prevent errors of omission such as this. One possible cause for the higher error rate with the UC-CD may have been unfamiliarity with the document. Only one subject indicated that he had any experience with the UC-CD and that he had only seen it but not used it. The average experience level for the UC-CD was 0.083 on a scale from 0 to 5 compared with 1.7 for the PAPER document. As noted earlier, the RP-WEB document was very similar to the PAPER document since it was developed from the PAPER document. The similarity between these two documents may have resulted in fewer errors as the subjects became familiar with the document during the testing. The reduction in errors in the second document may be evidence of this.

4) *The redesigned METOC images in the UC-CD did not decrease interpretation time but may have helped reduce browse time.*

A long line of research has shown that reading from computers is slower than reading from printed pages (e.g., Gould et al. 1987). The UC-CD is designed to make the images readable when displayed on computer in order to facilitate use. Computer fonts were used and font size was increased to be more compatible with low-resolution screens and color was added to make the UC-CD more interesting and usable. These changes did not produce measurable decreases in interpretation time over the RP-WEB document, but may have contributed to the decrease in navigation time with the UC-CD as compared with the RP-WEB. By using screen fonts and color for UC-CD, more information could be displayed in the same size image than was possible in the scanned images used in the RP-WEB document. This allowed the maps in UC-CD to show the entire region in one image and may have resulted in navigating to the information more quickly as compared to the RP-WEB document.

5) *The redesigned METOC images in the UC-CD have greater appeal than the scanned images of the RP-WEB document.*

The effort spent on redesigning the METOC images for computer and making them presentation quality was not wasted even if it did not produce measurable increases in efficiency, because it probably is responsible for the overwhelming enthusiasm for the UC-CD over the other documents. There were a greater number of images used in the briefs produced from the UC-CD than the other two documents, and the UC-CD was chosen for producing the best brief and being easiest to use. All 12 of the subjects felt that the UC-CD produced the best brief when asked: “Which document produced the best brief?” Nine out of 11 picked the RP-WEB document as their second choice. When asked: “Which document was easiest to use?” 10 of 12 subjects picked the UC-CD. Of the two who didn’t pick the UC-CD, one picked the RP-WEB but said he would have picked the UC-CD if he were used to it, and one picked the PAPER document because “I was familiar with it.” They both indicated that the UC-CD was their second choice. This is quite a contrast to Landauer’s (1995) findings; the subjects here preferred the digital documents over the PAPER document.

The enthusiasm for the UC-CD is understandable given that the METOC officer needs a high quality briefing. Discussions with the officers during the exit interview indicated that more and more frequently the officers prepare their briefs on computers and present them directly from the computer to an overhead display. Only when such a display is not available will they print to transparency. Given the dependence on the computer and its display, the efforts toward making graphics presentation quality as in the UC-CD will decrease the burden on the METOC officer.

6) The Web was observed to be somewhat unreliable.

Numerous delays were encountered in conducting this experiment because the Web or the server sites were down, sometimes for a few minutes, sometimes for hours. An extra trip to the second site had to be made to finish the testing because the Web went down. Although improvements in reliability are continually being made, any design using the Web will need to address its reliability.

5.2 Recommendations

The question is not so much which document to use but which interface strategies to incorporate. The media used will be determined by the tradeoffs arising from its physical properties. The trend is clearly toward computers and more specifically toward the Web. Regardless of the media, if the goal is to provide easy to use and efficient documents, we recommend the following:

1) Adopt the navigational strategies used in the UC-CD to speed navigation. This may be accomplished in Web documents through the use of frames.

2) Make the METOC officers aware of the electronic information that is available from NAVOCEANO for their use. Almost none of the officers were aware of the UC-CD or the RP-WEB documents, even those that had spent time in the Arabian Gulf.

3) Provide CDROMs of the Web documents to the METOC officers or encourage them to download the documents prior to deployment. The latest version of Microsoft’s Internet Explorer has an option to save entire Web sites locally for offline browsing. This would allow a METOC officer to download the latest versions prior to deployment and carry them with him without needing to rely on a Web connection. This is currently being done by NAVOCEANO.

4) Redesign METOC material for the media. This would include both preparing presentation quality graphics for the computer interface resulting in higher quality presentations and using the computational capability of the computer to create more intuitive interfaces as was done with the daylight interface in UC-CD.

5) Follow the steps outlined in Section 3 of this document and the user-centered design approach in Section 3.3.

6) Design the interfaces to prompt or mandate user input where necessary to reduce errors due to omissions.

6. ACKNOWLEDGMENTS

This work was supported by the U.S. Office of Naval Research. Our thanks to CDR Chris Gunderson, LCDR Bill Nisley II, and LCDR Roy Ledesma for coordinating the subjects and facilities. We also are deeply indebted to Janet Stroup for preparing the final document.

REFERENCES

Belloti, V. and Y. Rogers, 1997. "From Web Press to Web Pressure: Multimedia Representations and Multimedia Publishing," in *Proceedings of CHI 1997* (Atlanta, GA, March 22-27, 1997) ACM, New York, pp. 279-286.

Borges, J., I. Morales, and N. Rodríguez, 1996. "Guidelines for Designing Usable World Wide Web Pages," in *Proceedings of CHI 1996* (Vancouver, British Columbia, Canada, March 22-27, 1996) ACM, New York, pp. 277-278.

Gould, J.D., L. Alfaro, V. Barnes, R. Finn, N. Grischkowsky, and A. Minuto, 1987. "Reading is Slower from CRT Displays than from Paper: Attempts to Isolate a Single-variable Explanation," *Human Factors* **29**, 269-300.

Landauer, T.K., 1995. *The Trouble with Computers* (MIT Press, Cambridge, MA).

Miyamoto, R.T., P.M. Hardisty, and M.W. Stoermer, 1991. "Transfer of Navy's Environmental Guides to CD-ROM," APL-UW Technical Memorandum TM 17-91, Applied Physics Laboratory, University of Washington, Seattle, WA.

Narayanan, N. and M. Hegarty, 1997. "Model-Based Design of Hypermedia Presentations," in *Proceedings of CHI 1997 Extended Abstracts* (Atlanta, GA, USA, March 22-27, 1997) ACM, New York, pp. 355-356.

Nielsen, J., 1997. "User Interface Design for the WWW," in *Proceedings of CHI 1997 Extended Abstracts: Tutorials* (Atlanta, GA, USA, March 22-27, 1997) ACM, New York.

Pliske, R., D. Klinger, R. Hutton, B. Crandall, B. Knight, and G. Klein, 1997. "Understanding Skilled Weather Forecasting: Implications for Training and the Design of Forecasting Tools," Technical Report AL/HR-CR-1997-0003, Armstrong Laboratory, Human Resources Directorate, Brooks Air Force Base, TX, 78235-5352.

Segal, L.D. and A.D. Andre, 1993. Activity Catalog Tool (A.C.T.), NASA Contractor Report CR-177634, NASA ARC, Moffett Field, CA.

Sutcliffe, A. and P. Faraday, 1994. "Designing Presentation in Multimedia Interfaces," in *Proceedings of CHI 1994* (Boston, MA, USA, April 24-28, 1994) ACM, New York, pp. 92-98.

Vora, P., M. Helander, and V. Shalin, 1994. "Evaluation of the Influence of Interface Styles and Multiple Access Paths in Hypertext," in *Proceedings of CHI 1994* (Boston, MA, USA, April 24-28, 1994) ACM, New York, pp. 323-329.

Appendix A

STATISTICAL TESTS

The statistical tests used included the *t* test for differences between means and the Analysis of Variance (ANOVA). For the *t* test, the larger the value, the greater the difference in the means. The *p* value is the probability the obtained *t* value could have occurred by chance under the hypothesis of no difference between the means. For the ANOVA, a summary of the significant effects are presented in Tables A1 through A5, together with the significant differences found in a post hoc analysis using the Tukey test. Each table summarizes the significant effects for a particular independent variable. The F ratio in these tables is the variance due to the source effect divided by the variance of the residual (for example, the variance due to different media types divided by the variance due to different subjects). The larger the F ratio, the greater the effect the source variable had in relation to the variance in the error residual, and the more likely that the result is significant. The *p* value is the probability that the obtained F ratio could have occurred by chance, and is not caused by the source effect. Analysis of variance is conservative and is based on disproving a null hypothesis that there is no difference in the source effects (that there is no difference between paper, Web and CDROM, for example). The *p* value is the probability of rejecting the null hypothesis (declaring difference between media type) when, in fact, it is true (no difference between media). Naturally, the smaller the error, the better. For this study, results with error less than 5% are reported to be significant and results with error less than 10% are reported to show a tendency toward significance. This probability is typical for studies involving human subjects such as this one. Note that failure to reject the null hypothesis does not necessarily imply that it should be accepted. Failure to find a significant difference between media, for example, does not imply that no difference exists, only that it was not significantly different in this study. The difference may be swamped by a much larger difference from some unaccounted factor. The variance for this unaccounted factor would be pooled in the error residual that makes up the denominator of the F ratio, resulting in a small F ratio and a corresponding large *p*. The post hoc tests shown in the tables are used to determine which of the treatments (which media type, for example) are significantly different from the others.

Table A1— ANOVA Summary of Significant Document Effects on Browse Time

Problem	**Task**	**F Ratio**	***p***	**Post Hoc Tests**
One (Briefing)	Browse time	5.62	.01	RP-WEB>[UC-CD, PAPER]
Two (Tidal)	Browse time	3.22	.07	RP-WEB>PAPER @ p=.10
Three (Fishing)	Browse time	3.77	.05	UC-CD>PAPER

Note: on Problem 1, one subject took a long time to search for bathymetry using the RP-WEB

Table A2 — ANOVA Summary of Significant Document Effects on Other Task Times

Problem	**Task**	**F Ratio**	***p***	**Post Hoc Tests**
Two (Tidal)	Total time	7.05	<.00	RP-WEB>PAPER
Two (Tidal)	Total time	7.24	<.00	RP-WEB @ Site 1 > [all others]
Five (Daylight)	Read time	5.64	.01	PAPER>UC-CD

Table A3 — ANOVA Summary of Significant Site Effect

Problem	**Task**	**F Ratio**	***p***	**Post Hoc Tests**
Three	Browse time	6.81	.03	Site 1> Site 2

Table A4 — ANOVA Summary of Significant Effects Due to Order

Problem	**Task/dependent**	**F Ratio**	***p***	**Post Hoc Tests**
Brief prep	Total time	10.36	<.001	First>[second, third]
Brief prep	Search time	7.29	<.001	First>[second, third]
Brief prep	Produce time	9.31	<.001	First>[second, third]
Two	Total time	3.70	.05	None
Three	Total time	5.94	.01	First>[second, third]
Three	Read time	7.60	<.00	First>[second, third]
Four	Read time	11.77	<.00	First>[Second, Third]
Five	Total time	4.05	.04	First>Third
Five	Read time	4.26	.03	First>Third
Prepare brief	Total task freq	5.29	<.01	First>Second>Third
Prepare brief	Search freq	5.70	.01	First>Second>Third
Prepare brief	Generate freq	6.07	.02	First>Second=Third

Table A5 — ANOVA Summary of Effects on Error

Variable	**F Ratio**	***p***	**Post Hoc Tests**
Document	3.48	.055	UC-CD > [RP-WEB, PAPER] @ p = .087
Order	3.61	.051	First > Second @ p = .054

Appendix B

BRIEFING SCORECARD

Consent form signed	________________	
Questionnaire filled out	________________	
Introduction read	________________	
Media #1 Type	[DMARS / Paper / Web]	
Familiarization done	________________	
Brief location	________________	
#images	________________	
bathymetry map	[correct / incorrect / missing]	
characteristic tide	[correct / incorrect / missing]	
tide range	[correct / incorrect / missing]	
typical tide curve	[correct / incorrect / missing]	
wind direction spring	[correct / incorrect / missing]	
wind speed spring	[correct / incorrect / missing]	
wind direction fall	[correct / incorrect / missing]	
wind speed fall	[correct / incorrect / missing]	
Timing location	________________	
tide range	________________	[correct / incorrect]
fishing	[yes / no]	[correct / incorrect]
sound speed	________________	[correct / incorrect]
daylight	________________	[correct / incorrect]
Media #2 Type	[DMARS / Paper / Web]	
Familiarization done	________________	
Brief location	________________	
#images	________________	
bathymetry map	[correct / incorrect / missing]	
characteristic tide	[correct / incorrect / missing]	
tide range	[correct / incorrect / missing]	
typical tide curve	[correct / incorrect / missing]	
wind direction spring	[correct / incorrect / missing]	
wind speed spring	[correct / incorrect / missing]	
wind direction fall	[correct / incorrect / missing]	
wind speed fall	[correct / incorrect / missing]	

Timing location	________________	
tide range	________________	[correct / incorrect]
fishing	[yes / no]	[correct / incorrect]
sound speed	________________	[correct / incorrect]
daylight	________________	[correct / incorrect]
Media #3 Type	[DMARS / Paper / Web]	
Familiarization done	________________	
Brief location	________________	
#images	________________	
bathymetry map	[correct / incorrect / missing]	
characteristic tide	[correct / incorrect / missing]	
tide range	[correct / incorrect / missing]	
typical tide curve	[correct / incorrect / missing]	
wind direction spring	[correct / incorrect / missing]	
wind speed spring	[correct / incorrect / missing]	
wind direction fall	[correct / incorrect / missing]	
wind speed fall	[correct / incorrect / missing]	
Timing location	________________	
tide range	________________	[correct / incorrect]
fishing	[yes / no]	[correct / incorrect]
sound speed	________________	[correct / incorrect]
daylight	________________	[correct / incorrect]
Prepare image for paper document	________________	
Exit interview		
Which media was easiest to use?	[DMARS / Paper / Web]	
Which media produced the best brief?	[DMARS / Paper / Web]	
Please do not discuss until next week.		